

Weekly report (2012.5.6~13)

这周主要找一些论文看，了解并行、大规模绘制方面的工作（发现找文章好难）。另外，对 hadoop 作了一些了解，为之后将在 hadoop 上做的一些实验作准备。

文献阅读

[1] 《A Map Reduce Framework for Programming Graphics Processors》

- 会议：STMCS 08
(<http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=2276©ownerid=421>)
- 作者：Bryan Catanzaro, Narayanan Sundaram and Kurt Keutzer
- 这篇文章主要：
 - 介绍了他们在 GPU 上实现的一个 MapReduce 框架，是基于单个 GPU 实现的。
 - 这一框架允许用户自定义一个 map 函数，多个 reduce 操作，以及一个 cleanup 函数，然后会有一个代码生成器将用户编写的代码进行处理生成适合于 GPU 上运行的代码，以降低 GPU 编程的代价，同时保证 GPU 的利用率，提高效率。
 - 通过对 map 和部分 reduce 操作进行合并（本地 reduce）以及使用一些策略减少 communication 和 synchronization，作者表示相对于之前的工作，这种框架可以减少额外开销。
 - 在 reduce 阶段，进行循环展开，并提供了几种可选的方案组合让用户选择，以在不同的数据规模上最大程度地提高 reduce 阶段的并行性。
- 不足：
 - 数据流和 CPU 上的 MapReduce 类似，但是在程序编写上还是会涉及到 GPU 编程。
 - 在 map 和 reduce 阶段基于 GPU 做了相当多的假设，甚至假设 map 阶段只产生小于 32 个 key，也就是说这种 MapReduce 框架的通用性并不好。
- 个人觉得这篇文章的工作还是很 NB 的，从表面上看对单 GPU 的编程应该还是有较大意义的，但是对于我目前做的好像帮助不大。
- 文章链接：http://www.cs.berkeley.edu/~narayans/Publications_files/stmcs2008.pdf

[2] 《Above the Clouds: A Berkeley View of Cloud Computing》

- 时间：February 10, 2009
- 文章对云计算的发展做了较详尽的介绍，包括云计算的背景、现状、适用范围、发展障碍、机遇等等各方面，算是对云计算作了进一步的了解吧。
- 文章链接：<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>

下周计划

1. 学车场考。
2. 抽时间将现有的分布式绘制代码迁移到 hadoop 上，做一些实验。